

ON A NOVEL INTEGRABLE GENERALIZATION OF THE SINE-GORDON EQUATION

J. Lenells^a and A. S. Fokas^b

^aInstitut für Angewandte Mathematik, Leibniz Universität Hannover
Welfengarten 1, 30167 Hannover, Germany
E-mail: lenells@ifam.uni-hannover.de

^bDepartment of Applied Mathematics and Theoretical Physics, University of Cambridge,
Cambridge CB3 0WA, United Kingdom
E-mail: t.fokas@damtp.cam.ac.uk

Abstract

We consider an integrable generalization of the sine-Gordon (sG) equation that was earlier derived by one of the authors using bi-Hamiltonian methods. This equation is related to the sG equation in the same way that the Camassa-Holm equation is related to the KdV equation. In this paper we: (a) Derive a Lax pair. (b) Use the Lax pair to solve the initial value problem on the line. (c) Analyze solitons. (d) Show that the generalized sG and sG equations are related by a Liouville transformation. (e) Derive conservation laws. (f) Analyze traveling-wave solutions.

AMS SUBJECT CLASSIFICATION (2000): 35Q55, 37K15.

KEYWORDS: Integrable system, inverse spectral theory, Riemann-Hilbert problem, solitons.

1 Introduction

We consider the following integrable generalization of the sine-Gordon equation, which was derived in [6] using bi-Hamiltonian methods:

$$(1.1) \quad u_{tx} = (1 + \nu \partial_x^2) \sin(u), \quad x \in \mathbb{R}, \quad t > 0,$$

where $u(x, t)$ is a scalar-valued function and $\nu \in \mathbb{R}$ is a parameter. Equation (1.1) is related to the sine-Gordon equation (i.e. to the equation obtained from (1.1) by letting $\nu = 0$) in the same way that the Camassa-Holm (CH) equation (see [3, 11]) is related to the Korteweg-de Vries (KdV) equation. Actually, there exist even deeper analogies between (1.1) and the CH equation. Indeed, recall that for a particular class of initial conditions it is possible to use a Liouville transformation to map CH to an equation in the KdV hierarchy. Similarly, it will be shown here that for $\nu < 0$, equation (1.1) is related to the sine-Gordon equation via an appropriate Liouville type transformation (see Proposition 5.1). Furthermore, one of the distinguished features of the CH equation is that it possesses certain non-smooth traveling-wave solutions called peakons. Similarly, it will be shown here that equation (1.1), in addition to

smooth traveling-wave solutions (see Proposition 7.1), also possesses certain non-smooth traveling-wave solutions called cuspons (see Proposition 7.2).

Taking into consideration the extensive interest in the literature for the CH equation, as well as the above similarities between (1.1) and CH, it is natural to expect that equation (1.1) may also be of some interest.

This paper is organized as follows: Section 2 presents an algorithmic derivation of a Lax pair for equation (1.1), starting from the knowledge of the recursion operator (this approach might be useful also for other integrable equations generated via bi-Hamiltonian methods, since for such equations the associated recursion operator is known). Section 3 presents the solution of the initial-value problem of equation (1.1) with decaying initial conditions in the case that $\nu < 0$. The more interesting case of $\nu > 0$ remains open. Solitons and conservation laws are discussed in sections 4 and 6. The explicit Liouville transformation relating (1.1) with the sine-Gordon equation is presented in section 5. Smooth, as well as non-smooth traveling-wave solutions are derived in section 7.

Taking into consideration that (1.1) is related to the sG equation with an explicit transformation (for $\nu < 0$), it follows that section 3 also presents the solution of the initial-value problem of the sG equation. The solution of this problem was first presented in the classical work [1], however the method presented in section 3 is based on the ‘modern’ Riemann-Hilbert approach and is also influenced by the method of [7] (so that it is not necessary to formulate separately the time evolution of the spectral functions, see [10]).

After this work was completed, the references [2, 16] (where this equation is studied in connection with pseudospherical surfaces) and [17] (where the transformation between (1.1) and the sine-Gordon equation is presented) were brought to the authors attention.

2 A Lax pair

Let

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Equation (1.1) is the compatibility condition of the following Lax pair:

$$(2.1) \quad \begin{cases} \psi_x + i\lambda\sigma_3\psi = W_1\psi, \\ \psi_t + i\left(\nu\lambda - \frac{1}{4\lambda}\right)\sigma_3\psi = W_2\psi, \end{cases}$$

where $\psi(x, t, \lambda)$ is a 2×2 -matrix valued function, $\lambda \in \hat{\mathbb{C}} = \mathbb{C} \cup \infty$ is a spectral parameter, and $W_j(x, t, \lambda)$, $j = 1, 2$, are defined by

$$(2.2) \quad \begin{aligned} W_1 &= u_x \left(\frac{i}{2}\sigma_1 + \lambda\sqrt{\nu}\sigma_2 \right), \\ W_2 &= \frac{i}{4\lambda} [\sin(u)\sigma_2 + (\cos(u) - 1)\sigma_3] + \left[\frac{i\nu}{2} \cos(u)u_x - \frac{\sqrt{\nu}}{2} \sin(u) \right] \sigma_1 \\ &\quad + \lambda \left[\nu^{3/2} \cos(u)u_x\sigma_2 - i\nu(\cos(u) - 1)\sigma_3 \right]. \end{aligned}$$

2.1 Derivation of Lax pair

Before deriving the Lax pair (2.1), we first recall the case of the sine-Gordon equation

$$(2.3) \quad q_t = \sin(\partial_x^{-1} q).$$

Equation (2.3) admits the isospectral problem (x -part of a Lax pair)

$$(2.4) \quad \psi_x = L\psi, \quad L = -i\lambda\sigma_3 + \frac{i}{2}q\sigma_1,$$

where $\psi(x, \lambda)$ is a vector valued eigenfunction and $\lambda \in \mathbb{C}$ is a spectral parameter, as well as the recursion operator R defined by

$$R = \theta_2 \partial_x^{-1}, \quad \theta_2 = \partial_x^3 + \partial_x q \partial_x^{-1} q \partial_x.$$

Differentiating equation (2.4) with respect to q in the direction w (where ψ and λ are considered functionals of q), we find

$$(2.5) \quad \partial_x \psi'(q)[w] = L\psi'(q)[w] - i \int G_\lambda w dx \sigma_3 \psi + \frac{i}{2} w \sigma_1 \psi,$$

where G_λ denotes the gradient of λ defined by

$$\lambda'(q)[w] = \int G_\lambda w dx \quad \text{for all } w.$$

Let ψ^* be a solution of the adjoint equation

$$(2.6) \quad \psi_x^* = -L^A \psi^*,$$

where $L^A = L$ denotes the adjoint of L with respect to the inner product

$$(f, g) = \int f^T g dx, \quad \text{i.e.} \quad (f, Lg) = (L^A f, g).$$

Multiplying (2.5) by ψ^{*T} from the left and integrating the resulting equation with respect to x , we find

$$(2.7) \quad \int \psi^{*T} \partial_x \psi'(u)[w] dx = \int \psi^{*T} L\psi'(u)[w] dx \\ + \int \left[G_\lambda \left(-i \int \psi^{*T} \sigma_3 \psi dx \right) + \frac{i}{2} \psi^{*T} \sigma_1 \psi \right] w dx.$$

The first two terms of (2.7) cancel in view of (2.6). Thus, since w is arbitrary,

$$(2.8) \quad G_\lambda = \frac{\psi^{*T} \sigma_1 \psi}{2 \int \psi^{*T} \sigma_3 \psi dx}.$$

From the bi-Hamiltonian theory, we expect the gradient G_λ to satisfy an eigenvalue equation of the form [8]

$$(2.9) \quad R^A G_\lambda = \mu(\lambda) G_\lambda,$$

where $\mu = \mu(\lambda)$ is an eigenvalue and $R^A = \partial_x^{-1}\theta_2$ denotes the adjoint of R . In fact, letting $\varphi = \psi^{*T}\sigma_1\psi$, we find that φ satisfies the equation

$$\left(\frac{\mu(\lambda)\varphi - \varphi_{xx}}{q}\right)_x = q\varphi_x, \quad \mu(\lambda) = -4\lambda^2,$$

which simplifies to (2.9) (up to the irrelevant constant in the denominator of (2.8)).

Equation (2.9) shows how the eigenfunctions of the x -part are related to the recursion operator for the sG equation. The eigenfunction $\varphi = \psi^{*T}\sigma_1\psi$ is sometimes referred to as a ‘squared eigenfunction’ since it is bilinear in ψ^* and ψ .

We now proceed to the derivation of the Lax pair (2.1). Equation (1.1) possesses the recursion operator [6]

$$R = \theta_2\theta_1^{-1},$$

where

$$\theta_1 = \partial_x + \nu\partial_x^3, \quad \theta_2 = \partial_x^3 + \partial_x u_x \partial_x^{-1} u_x \partial_x.$$

In order to find a Lax pair of the generalized sine-Gordon equation (1.1) we have to trace the steps used for the derivation of (2.9) backwards. Namely, knowing the recursion operator $R = \theta_2\theta_1^{-1}$, the goal is to determine the corresponding x -part. Thus, we seek an x -part such that the gradient G_λ of the spectral parameter λ of this x -part satisfies (2.9) with $R^A = \theta_1^{-1}\theta_2$. We let $q = u_x$ and make the ansatz

$$\psi_x = L\psi, \quad L = -i\lambda\sigma_3 + f(q, \lambda)\sigma_1 + g(q, \lambda)\sigma_2,$$

where f and g are scalar-valued functionals of q and of the x -derivatives of q which depend on λ . The analog of equation (2.5) is

$$(2.10) \quad \begin{aligned} \partial_x \psi'(q)[w] = & L\psi'(q)[w] - i \int G_\lambda w dx \sigma_3 \psi + \frac{\partial f}{\partial \lambda} \int G_\lambda w dx \sigma_1 \psi \\ & + \frac{\partial g}{\partial \lambda} \int G_\lambda w dx \sigma_2 \psi + (D_f w) \sigma_1 \psi + (D_g w) \sigma_2 \psi, \end{aligned}$$

where the operators D_f and D_g are defined by

$$D_f = \frac{\partial f}{\partial q} + \frac{\partial f}{\partial q_x} \partial_x + \frac{\partial f}{\partial q_{xx}} \partial_x^2 + \dots, \quad D_g = \frac{\partial g}{\partial q} + \frac{\partial g}{\partial q_x} \partial_x + \frac{\partial g}{\partial q_{xx}} \partial_x^2 + \dots.$$

Letting ψ^* be a solution of (2.6), multiplying (2.10) by ψ^{*T} from the left and integrating the resulting equation with respect to x , we find the following equation:

$$(2.11) \quad \begin{aligned} \int \psi^{*T} \partial_x \psi'(u)[w] dx = & \int \psi^{*T} L\psi'(u)[w] dx \\ & + \int \left[G_\lambda \left(-i \int \psi^{*T} \sigma_3 \psi dx + \int \psi^{*T} \frac{\partial f}{\partial \lambda} \sigma_1 \psi dx + \int \psi^{*T} \frac{\partial g}{\partial \lambda} \sigma_2 \psi dx \right) \right. \\ & \left. + D_f^A(\psi^{*T} \sigma_1 \psi) + D_g^A(\psi^{*T} \sigma_2 \psi) \right] w dx. \end{aligned}$$

The first two terms of this equation cancel and hence G_λ is given by

$$G_\lambda = -\frac{D_f^A(\psi^{*T}\sigma_1\psi) + D_g^A(\psi^{*T}\sigma_2\psi)}{-i\int\psi^{*T}\sigma_3\psi dx + \int\psi^{*T}\frac{\partial f}{\partial\lambda}\sigma_1\psi dx + \int\psi^{*T}\frac{\partial g}{\partial\lambda}\sigma_2\psi dx}.$$

Letting

$$f(q, \lambda) = a(\lambda)q, \quad g(q, \lambda) = b(\lambda)q,$$

where $a(\lambda)$ and $b(\lambda)$ are complex-valued functions of λ , we find that the function φ defined by

$$\varphi := -D_f^A(\psi^{*T}\sigma_1\psi) - D_g^A(\psi^{*T}\sigma_2\psi),$$

satisfies the equation

$$(2.12) \quad \left(\frac{\mu(\lambda)(\varphi + \nu\varphi_{xx}) - \varphi_{xx}}{q} \right)_x = q\varphi_x,$$

whenever

$$(2.13) \quad \mu(\lambda) = -\frac{4\lambda^2}{1 - 4\nu\lambda^2}, \quad a = \frac{i}{2}\sqrt{1 + 4b^2 - 4\lambda^2\nu}.$$

Rearranging (2.12), we deduce that G_λ satisfies (2.9) whenever μ, a, b satisfy (2.13). Choosing for simplicity $a = i/2$ and $b = \lambda\sqrt{\nu}$, we find

$$\psi_x = L\psi, \quad L = -i\lambda\sigma_3 + \frac{i}{2}q\sigma_1 + \lambda\sqrt{\nu}q\sigma_2.$$

Having obtained the x -part of the Lax pair (2.1), in order to find the corresponding t -part, we make the ansatz

$$\psi_t = M\psi, \quad M = \sum_{j=-1}^1 \lambda^j (A_j\sigma_1 + B_j\sigma_2 + C_j\sigma_3),$$

where A_j, B_j, C_j are scalar-valued functionals of q independent of λ . Identifying terms of $O(\lambda^j)$, $j = -1, 0, 1, 2$, in the compatibility equation $L_t - M_x + [L, M] = 0$ and using that $q_t = (1 + \nu\partial_x^2)\sin(\partial_x^{-1}q)$, long but straightforward computations lead to the t -part in (2.1).

3 Spectral analysis on the line

The Riemann-Hilbert (RH) formalism for integrating a nonlinear evolution equation is based on the construction of eigenfunctions of the associated Lax pair which can be joined together to a bounded and sectionally analytic function on the Riemann sphere of the spectral parameter $\lambda \in \hat{\mathbb{C}}$. In order to define eigenfunctions which possess the asymptotics

$$\mu = I + O(1/\lambda), \quad \lambda \rightarrow \infty,$$

we will use a different representation of the Lax pair (2.1). This representation involves an eigenfunction ϕ , which is related to ψ via the gauge transformation $\psi =$

$g\phi$. The form of g is such that it (a) diagonalizes the highest order terms in λ as $\lambda \rightarrow \infty$ and (b) preserves the natural symmetry properties of the Lax pair.

The spectral analysis of the Lax pair satisfied by ϕ is standard: We introduce two particular eigenfunctions via integration from $x = \pm\infty$, and use these eigenfunctions to formulate a 2×2 -matrix RH problem with jump across the real axis. The jump matrix is given in terms of two spectral functions $a(\lambda)$ and $b(\lambda)$, which are defined in terms of the initial data $u_0(x) = u(x, 0)$. The possible poles of $a(\lambda)$ give rise to singularities of the RH problem which correspond to solitons.

The solution $u(x, t)$ can be recovered from the solution of the RH problem, so that this approach provides the solution of the Cauchy problem for equation (1.1). However, the RH problem is naturally formulated in terms of variables (y, t) rather than (x, t) where y is a new variable. Therefore, we will only obtain the solution $u(x, t)$ in parametric form. This type of parametric representation arises also in the analysis of the Camassa-Holm and Degasperis-Procesi equations (cf. [4, 12] in the case of the Camassa-Holm equation).

In the rest of the paper we assume that $\nu < 0$ and then for simplicity let $\nu = -1$.

3.1 Symmetries

Assuming that $\nu = -1$, the Lax pair (2.1) can be written as

$$(3.1) \quad \begin{cases} \psi_x = L\psi, \\ \psi_t = M\psi, \end{cases}$$

where

$$L = -i\lambda\sigma_3 + W_1, \quad M = i\left(\lambda + \frac{1}{4\lambda}\right)\sigma_3 + W_2,$$

and W_1, W_2 are given by (2.2) with $\sqrt{\nu}$ replaced by i . Since L and M are trace-less and satisfy the symmetry relations

$$L^\dagger(\bar{\lambda}) = -L(\lambda), \quad M^\dagger(\bar{\lambda}) = -M(\lambda),$$

where A^\dagger denotes the complex-conjugate transpose of a matrix A , it follows that the functions

$$\det(\psi(x, t, \lambda)) \quad \text{and} \quad \psi^\dagger(x, t, \bar{\lambda})\psi(x, t, \lambda)$$

are independent of x and t . We deduce that ψ satisfies

$$(3.2) \quad \det(\psi(x, t, \lambda)) = 1, \quad \psi^\dagger(x, t, \bar{\lambda}) = \psi^{-1}(x, t, \lambda),$$

for all (x, t) provided that these equalities hold at only one point (x_0, t_0) . The eigenfunctions introduced in the sequel will indeed have this property.

3.2 Eigenfunctions

Suppose that $u(x, t)$ is a sufficiently smooth real-valued solution of (1.1) and suppose that $\cos(u(x, t)) - 1$ has decay as $x \rightarrow \pm\infty$ for all $t \geq 0$. Then the form of equation (1.1) implies that $\int_{\mathbb{R}} \sin(u(x, t))dx = 0$ for each t . Define $m(x, t)$ by

$$(3.3) \quad m(x, t) = 1 + u_x^2(x, t).$$

The gauge transformation

$$(3.4) \quad \psi(x, t, \lambda) = g(x, t) \phi(x, t, \lambda),$$

where

$$(3.5) \quad g(x, t) = \sqrt{\frac{1 + \sqrt{m}}{2\sqrt{m}}} \begin{pmatrix} 1 & i \frac{1 - \sqrt{m}}{u_x} \\ i \frac{1 - \sqrt{m}}{u_x} & 1 \end{pmatrix},$$

transforms the Lax pair (2.1) into

$$(3.6) \quad \begin{cases} \phi_x + i\lambda p_x \sigma_3 \phi = V_1 \phi, \\ \phi_t + i(\lambda p_t - \frac{1}{4\lambda}) \sigma_3 \phi = V_2 \phi, \end{cases}$$

where $V_1(x, t)$ and $V_2(x, t, \lambda)$ are defined by

$$\begin{aligned} V_1 &= \frac{i}{2} \left(u_x + \frac{u_{xx}}{m} \right) \sigma_1, \\ V_2 &= -\frac{i}{4\lambda} \sigma_3 + i \frac{\cos(u) - u_x \sin(u)}{4\sqrt{m}\lambda} \sigma_3 + i \frac{u_x \cos(u) + \sin(u)}{4\sqrt{m}\lambda} \sigma_2 \\ &\quad - \frac{i}{2} \left(u_x + \frac{u_{xx}}{m} \right) \cos(u) \sigma_1 \end{aligned}$$

and $p(x, t)$ is a real-valued function defined by

$$(3.7) \quad p(x, t) = x - t + \int_{-\infty}^x (\sqrt{m(x', t)} - 1) dx'.$$

The conservation law

$$(3.8) \quad (\sqrt{m})_t + (\cos(u)\sqrt{m})_x = 0$$

implies that $p_t = -\cos(u)\sqrt{m}$. Note that definition (3.3) of m implies that g is free of singularities also at points where $u_x = 0$.

The form (3.5) of g is motivated by the fact that it diagonalizes the terms of $O(\lambda)$ of the Lax pair (2.1) and that it satisfies

$$(3.9) \quad \det(g(x, t)) = 1, \quad g^\dagger(x, t) = g^{-1}(x, t).$$

The relations (3.9) ensure that the gauge transformation (3.4) preserves the two properties in (3.2), i.e.

$$(3.10) \quad \det(\phi(x, t, \lambda)) = 1, \quad \phi^\dagger(x, t, \bar{\lambda}) = \phi^{-1}(x, t, \lambda).$$

We have defined V_1 and V_2 so that

$$V_1(x, t) \rightarrow 0, \quad V_2(x, t, \lambda) \rightarrow 0, \quad x \rightarrow \pm\infty.$$

The form of the Lax pair (3.6) is convenient for the definition of eigenfunctions which are well-behaved near $\lambda = \infty$. Introducing an eigenfunction μ by

$$\phi = \mu e^{-i(\lambda p - \frac{t}{4\lambda})\sigma_3},$$

we find that the Lax pair (3.6) can be written as

$$(3.11) \quad \begin{cases} \mu_x + i\lambda p_x[\sigma_3, \mu] = V_1\mu, \\ \mu_t + i(\lambda p_t - \frac{1}{4\lambda})[\sigma_3, \mu] = V_2\mu. \end{cases}$$

We define two eigenfunctions μ_{\pm} of (3.11) as the solutions of the following two Volterra integral equations

$$(3.12) \quad \begin{aligned} \mu_+(x, t, \lambda) &= I - \int_x^\infty e^{i\lambda(p(x', t) - p(x, t))\hat{\sigma}_3} V_1(x', t) \mu_+(x', t, \lambda) dx', \\ \mu_-(x, t, \lambda) &= I + \int_{-\infty}^x e^{i\lambda(p(x', t) - p(x, t))\hat{\sigma}_3} V_1(x', t) \mu_-(x', t, \lambda) dx', \end{aligned}$$

where $\hat{\sigma}_3$ acts on a 2×2 matrix A by $\hat{\sigma}_3 A = [\sigma_3, A]$. The second columns of these equations involve the exponential $e^{2i\lambda(p(x', t) - p(x, t))}$. Since $p(x, t)$ is an increasing function of x for any fixed t , we deduce that the second columns of μ_+ and μ_- are bounded and analytic for λ in the upper and lower halves of the complex λ -plane, respectively. We will use the superscripts $(+)$ and $(-)$ to denote these boundedness properties. Similar considerations are valid for the first columns and hence

$$\mu_+ = \left(\mu_+^{(-)}, \mu_+^{(+)} \right), \quad \mu_- = \left(\mu_-^{(+)}, \mu_-^{(-)} \right).$$

The Lax pair (3.11) is of ‘standard form’ as $\lambda \rightarrow \infty$, i.e. the highest-order terms of $O(\lambda)$ are diagonal and appear in the exponents of (3.12), whereas the next order terms of $O(1)$ are off-diagonal. Substitution into (3.11) of the expansion

$$\mu = D + \frac{\mu_1}{\lambda} + \frac{\mu_2}{\lambda^2} + \cdots, \quad \lambda \rightarrow \infty,$$

shows that D is a diagonal matrix independent of x and t . In particular, μ_+ and μ_- have the following asymptotics near $\lambda = \infty$:

$$\begin{aligned} \left(\mu_-^{(+)}, \mu_+^{(+)} \right) &= I + O(1/\lambda), \quad \lambda \rightarrow \infty, \quad \text{Im } \lambda \geq 0, \\ \left(\mu_+^{(-)}, \mu_-^{(-)} \right) &= I + O(1/\lambda), \quad \lambda \rightarrow \infty, \quad \text{Im } \lambda \leq 0. \end{aligned}$$

Note that μ_+ and μ_- are nonsingular near $\lambda = 0$, because the integral equations (3.12) involve only the function V_1 and not V_2 .

3.3 Spectral functions

We define the spectral function $s(\lambda)$ by

$$(3.13) \quad \mu_+(x, t, \lambda) = \mu_-(x, t, \lambda) e^{-i(\lambda p(x, t) - \frac{t}{4\lambda})\hat{\sigma}_3} s(\lambda), \quad \text{Im } \lambda = 0.$$

Evaluation of (3.13) at $t = 0$ and $x \rightarrow -\infty$ gives

$$(3.14) \quad s(\lambda) = I - \int_{-\infty}^{\infty} e^{i\lambda p(x,0)\hat{\sigma}_3} V_1(x,0) \mu_+(x,0,\lambda) dx, \quad \text{Im } \lambda = 0.$$

We deduce from (3.10) that μ_{\pm} have the properties

$$(3.15) \quad \det(\mu_{\pm}(x,t,\lambda)) = 1, \quad \mu_{\pm}(x,t,\lambda)_{11} = \overline{\mu_{\pm}(x,t,\bar{\lambda})_{22}}, \quad \mu_{\pm}(x,t,\lambda)_{21} = -\overline{\mu_{\pm}(x,t,\bar{\lambda})_{12}}.$$

This implies that

$$\det s(\lambda) = 1$$

and that there exist functions $a(\lambda)$ and $b(\lambda)$ such that

$$s(\lambda) = \begin{pmatrix} \overline{a(\bar{\lambda})} & b(\lambda) \\ -\overline{b(\bar{\lambda})} & a(\lambda) \end{pmatrix}.$$

From the explicit expression (3.14) for $s(\lambda)$ and the fact that the second column of μ_+ is defined and analytic in $\text{Im } \lambda > 0$, we find that $a(\lambda)$ has an analytic continuation to the upper half-plane.

3.4 Residue conditions

We assume that $a(\lambda)$ has N simple zeros $\{\lambda_j\}_{j=1}^N$ in the upper half-plane. These eigenvalues are either purely imaginary or arise as complex conjugate pairs $(\lambda, -\bar{\lambda})$. The second column of equation (3.13) is

$$(3.16) \quad \mu_+^{(+)} = a\mu_-^{(-)} + b\mu_-^{(+)} e^{-2i(\lambda p - \frac{t}{4\lambda})}, \quad \text{Im } \lambda = 0.$$

Applying $\det(\mu_-^{(+)}, \cdot)$ to this equation and recalling (3.15), we find

$$\det \begin{pmatrix} \mu_-^{(+)}(x,t,\lambda) & \mu_+^{(+)}(x,t,\lambda) \end{pmatrix} = a(\lambda), \quad \text{Im } \lambda \geq 0,$$

where we have used that both sides are well-defined and analytic in the upper half-plane to extend the relation to $\text{Im } \lambda \geq 0$. Hence, if $a(\lambda_j) = 0$, then $\mu_-^{(+)}(x,t,\lambda_j)$ and $\mu_+^{(+)}(x,t,\lambda_j)$ are linearly dependent vectors for each x and t . It follows that there exist constants b_j such that

$$(3.17) \quad \mu_-^{(+)}(x,t,\lambda_j) = b_j e^{2i(\lambda_j p - \frac{t}{4\lambda_j})} \mu_+^{(+)}(x,t,\lambda_j), \quad x \in \mathbb{R}, t > 0.$$

Recalling the symmetries in (3.15), the complex conjugate of (3.17) is

$$\mu_-^{(-)}(x,t,\bar{\lambda}_j) = -\bar{b}_j e^{-2i(\bar{\lambda}_j p - \frac{t}{4\bar{\lambda}_j})} \mu_+^{(-)}(x,t,\bar{\lambda}_j), \quad x \in \mathbb{R}, t > 0.$$

Consequently, the residues of $\mu_-^{(+)}/a$ and $\mu_-^{(-)}/\bar{a}$ at λ_j and $\bar{\lambda}_j$ are

$$\begin{aligned}\operatorname{Res}_{\lambda_j} \frac{\mu_-^{(+)}(x, t, \lambda)}{a(\lambda)} &= C_j e^{2i(\lambda_j p(x, t) - \frac{t}{4\lambda_j})} \mu_+^{(+)}(x, t, \lambda_j), \\ \operatorname{Res}_{\bar{\lambda}_j} \frac{\mu_-^{(-)}(x, t, \lambda)}{a(\bar{\lambda})} &= -\bar{C}_j e^{-2i(\bar{\lambda}_j p(x, t) - \frac{t}{4\bar{\lambda}_j})} \mu_+^{(-)}(x, t, \bar{\lambda}_j),\end{aligned}$$

where $\dot{a} = \frac{da}{d\lambda}$ and $C_j = \frac{b_j}{\dot{a}(\lambda_j)}$.

3.5 Riemann-Hilbert problem

We now describe how the solution to the initial-value problem for equation (1.1) on the line can be expressed in terms of the solution of a 2×2 -matrix Riemann-Hilbert problem. Relation (3.13) can be rewritten in the form of the RH problem

$$M_-(x, t, \lambda) = M_+(x, t, \lambda)J(x, t, \lambda), \quad \operatorname{Im} \lambda = 0,$$

where the matrices M_- , M_+ , J are defined by

$$(3.18) \quad M_+ = \left(\frac{\mu_-^{(+)}}{a(\lambda)}, \mu_+^{(+)} \right), \quad \operatorname{Im} \lambda \geq 0;$$

$$M_- = \left(\mu_+^{(-)}, \frac{\mu_-^{(-)}}{a(\bar{\lambda})} \right), \quad \operatorname{Im} \lambda \leq 0;$$

$$(3.19) \quad J = \begin{pmatrix} 1 & -\frac{b(\lambda)}{a(\bar{\lambda})} e^{-2i(\lambda p - \frac{t}{4\lambda})} \\ -\frac{\bar{b}(\bar{\lambda})}{a(\lambda)} e^{2i(\lambda p - \frac{t}{4\lambda})} & \frac{1}{a(\lambda)a(\bar{\lambda})} \end{pmatrix}, \quad \operatorname{Im} \lambda = 0.$$

The contour for this RH problem is the real axis.

However, since the jump matrix J , in addition to the spectral functions $a(\lambda)$ and $b(\lambda)$, also depends on the function $p(x, t)$, this RH problem cannot be formulated in terms of the initial data alone. We overcome this problem by changing variables

$$(x, t) \mapsto (y, t), \quad y = p(x, t).$$

Since J only depends on x through $y = p(x, t)$ we can formulate a family of RH problem parametrized by (y, t) . From the solution $M(y, t, \lambda)$ of this RH problem, we recover $u(x, t)$ in parametric form.

Defining

$$(3.20) \quad J^{(y)}(y, t, \lambda) = \begin{pmatrix} 1 & -\frac{b(\lambda)}{a(\bar{\lambda})} e^{-2i(\lambda y - \frac{t}{4\lambda})} \\ -\frac{\bar{b}(\bar{\lambda})}{a(\lambda)} e^{2i(\lambda y - \frac{t}{4\lambda})} & \frac{1}{a(\lambda)a(\bar{\lambda})} \end{pmatrix}, \quad \operatorname{Im} \lambda = 0,$$

we have the following result.

Theorem 3.1 *Let $u_0(x)$, $x \in \mathbb{R}$, be such that $\cos(u_0(x)) - 1$ has sufficient smoothness and decay as $x \rightarrow \pm\infty$ (for details see (3.35) below) and such that $\int_{\mathbb{R}} \sin(u_0) dx = 0$. Define $V_{10}(x)$, $m_0(x)$, and $p_0(x)$ by*

$$V_{10} = \frac{i}{2} \left(u_{0x} + \frac{u_{0xx}}{m_0} \right) \sigma_1, \quad m_0 = 1 + u_{0x}^2, \quad p_0(x) = x + \int_{-\infty}^x (\sqrt{m_0(x')} - 1) dx'.$$

Let $\mu_+(x, 0, \lambda)$ and $\mu_-(x, 0, \lambda)$ be the unique solutions of the Volterra linear integral equations (3.12) evaluated at $t = 0$, with $V_1(x, 0) = V_{10}(x)$ and $p(x, 0) = p_0(x)$. Define $\{a(\lambda), b(\lambda), C_j\}$ by

$$(3.21) \quad \begin{pmatrix} b(\lambda) \\ a(\lambda) \end{pmatrix} = [s(\lambda)]_2, \quad s(\lambda) = I - \int_{-\infty}^{\infty} e^{i\lambda p_0(x)\hat{\sigma}_3} V_{10}(x) \mu_+(x, 0, \lambda) dx,$$

and

$$(3.22) \quad [\mu_-(x, 0, \lambda_j)]_1 = \dot{a}(\lambda_j) C_j e^{2i(\lambda_j p_0(x) - \frac{t}{4\lambda_j})} [\mu_+(x, 0, \lambda_j)]_2, \quad j = 1, \dots, N,$$

where $[A]_1$ and $[A]_2$ denote the first and second columns of a 2×2 -matrix A . We assume that $a(\lambda)$ has N simple zeros $\{\lambda_j\}_{j=1}^N$ in the upper half-plane.

Then

- $a(\lambda)$ is defined for $\text{Im } \lambda \geq 0$, analytic in $\text{Im } \lambda > 0$, and $a(\lambda) = 1 + O(1/\lambda)$, $\lambda \rightarrow \infty$.
- $b(\lambda)$ is defined for $\lambda \in \mathbb{R}$ and $b(\lambda) = O(1/\lambda)$, $\lambda \rightarrow \infty$.
- $|a(\lambda)|^2 + |b(\lambda)|^2 = 1$, $\lambda \in \mathbb{R}$.

Suppose there exists a unique solution $u(x, t)$ of equation (1.1) with initial data $u_0(x)$ such that $\cos(u(x, t)) - 1$ has sufficient smoothness and decay as $x \rightarrow \pm\infty$ for each $t \geq 0$. Then $u(x, t)$ is given in parametric form by

$$(3.23a) \quad u(X(y, t), t) = 2\text{Im} \left(\int_{-\infty}^y \alpha(y', t) dy' \right) + 2n\pi,$$

where $2n\pi = \lim_{x \rightarrow -\infty} u_0(x)$, the function $X(y, t)$ is defined by

$$(3.23b) \quad X(y, t) = y + t + \int_{-\infty}^y \left(\sqrt{1 - 4\text{Im}(\alpha(y', t))^2} - 1 \right) dy',$$

the function $\alpha(y, t)$ is the unique solution of the Ricatti equation

$$(3.23c) \quad \alpha_y = \alpha^2 - 4i \left(\lim_{\lambda \rightarrow \infty} \lambda M_{12}(y, t, \lambda) \right) \alpha - \frac{1}{4}, \quad \lim_{y \rightarrow \pm\infty} \alpha(y, t) = -\frac{1}{2},$$

which satisfies

$$|\alpha(y, t)| = \frac{1}{2}, \quad y \in \mathbb{R}, \quad t \geq 0,$$

and $M(y, t, \lambda)$ is the unique solution of the following RH problem:

- $M(y, t, \lambda) = \begin{cases} M_+(y, t, \lambda), & \text{Im } \lambda \geq 0, \\ M_-(y, t, \lambda), & \text{Im } \lambda \leq 0, \end{cases}$ is sectionally meromorphic.
- $M_-(y, t, \lambda) = M_+(y, t, \lambda)J^{(y)}(y, t, \lambda)$ for $\lambda \in \mathbb{R}$, where $J^{(y)}$ is defined in (3.20).
- $M(y, t, \lambda)$ has the asymptotic behavior

$$(3.24) \quad M(y, t, \lambda) = I + O\left(\frac{1}{\lambda}\right), \quad \lambda \rightarrow \infty.$$

- The first column of M_+ has simple poles at $\lambda = \lambda_j$, $j = 1, \dots, N$, and the second column of M_- has simple poles at $\lambda = \bar{\lambda}_j$, $j = 1, \dots, N$. The associated residues are given by

$$(3.25a) \quad \text{Res}_{\lambda_j}[M(y, t, \lambda)]_1 = C_j e^{2i(\lambda_j y - \frac{t}{4\lambda_j})} [M(y, t, \lambda_j)]_2,$$

$$(3.25b) \quad \text{Res}_{\bar{\lambda}_j}[M(y, t, \lambda)]_2 = -\bar{C}_j e^{-2i(\bar{\lambda}_j y - \frac{t}{4\bar{\lambda}_j})} [M(y, t, \bar{\lambda}_j)]_1.$$

Proof. It remains to prove (3.23). By substituting the expansion

$$\mu(x, t, \lambda) = I + \frac{\mu^{(1)}(x, t)}{\lambda} + \frac{\mu^{(2)}(x, t)}{\lambda^2} + O\left(\frac{1}{\lambda^3}\right), \quad \lambda \rightarrow \infty,$$

into the x -part of (3.11), we find by considering the terms of $O(1)$ that

$$(3.26) \quad 4\mu_{12}^{(1)}(x, t) = \frac{1}{\sqrt{m(x, t)}} \left(u_x(x, t) + \frac{u_{xx}(x, t)}{m(x, t)} \right).$$

Define $Q(y, t)$ by $Q_y(p(x, t), t) = -4\mu_{12}^{(1)}(x, t)$. Then, by construction of the RH problem,

$$(3.27) \quad Q_y(y, t) = -4 \lim_{\lambda \rightarrow \infty} (\lambda M_{12}(y, t, \lambda)).$$

Equation (3.26) is conveniently expressed in terms of $y = p(x, t)$ and Q . Indeed, using that $\partial_x = \sqrt{m}\partial_y$ and $m = 1/(1 - u_y^2)$, we find

$$(3.28) \quad Q_y = - \left(u_y + \frac{u_{yy}}{\sqrt{1 - u_y^2}} \right).$$

Integration of this equation yields

$$(3.29) \quad Q = -u - \arcsin(u_y),$$

where we have fixed the integration constant by assuming that

$$Q(y, t) + u(y, t) \rightarrow 0, \quad y \rightarrow \pm\infty.$$

Note that

$$-1 \leq u_y = \frac{u_x}{\sqrt{1+u_x^2}} \leq 1,$$

so that Q is well-defined by (3.29). We rewrite (3.29) as

$$(3.30) \quad u_y = -\sin(Q+u).$$

Defining $\alpha(y, t)$ by

$$(3.31) \quad \alpha = -\frac{1}{2}e^{i(Q+u)},$$

equation (3.30) becomes the Ricatti equation

$$(3.32) \quad \alpha_y = \alpha^2 + iQ_y\alpha - \frac{1}{4}, \quad \alpha(\pm\infty, t) = -\frac{1}{2}.$$

By (3.30) and (3.31),

$$(3.33) \quad u_y = 2\text{Im}(\alpha).$$

On the other hand, $y \rightarrow x - t$ as $x \rightarrow -\infty$, and so

$$(3.34) \quad x = y + t + \int_{-\infty}^y \left(\sqrt{1 - u_y^2} - 1 \right) dy'.$$

The parametric formula (3.23) now follows from equations (3.27), (3.32), (3.33), and (3.34). \square

The jump matrix $J^{(y)}$ defined in (3.20) satisfies the symmetry relation $\overline{J_{12}^{(y)}(\bar{\lambda})} = J_{21}^{(y)}(\lambda)$. This implies that there exists a ‘vanishing lemma’ for the RH problem. This guarantees existence in $H_1(\mathbb{R})$ [9]. Thus the initial conditions must be chosen in an appropriate function space such that

$$(3.35) \quad p_0(x) - x \in L_\infty, \quad V_{10} \in L_1, \quad a, b \in H_1.$$

4 Soliton solutions

The solitons correspond to spectral data $\{a(\lambda), b(\lambda), C_j\}$ for which $b(\lambda)$ vanishes identically. In this case the jump matrix $J^{(y)}$ in (3.20) is the identity matrix and the RH problem of Theorem 3.1 consists of finding a meromorphic function $M(y, t, \lambda)$ satisfying (3.24) and the residue conditions (3.25). In what follows we will suppress the (y, t) -dependence of M and write $M(\lambda)$ for $M(y, t, \lambda)$.

Let $\theta_j = \lambda_j y - \frac{t}{4\lambda_j}$. From (3.24) and (3.25a) we find

$$(4.1) \quad [M(\lambda)]_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j=1}^N \frac{C_j e^{2i\theta_j} [M(\lambda_j)]_2}{\lambda - \lambda_j}.$$

It follows from (3.15) that $M(\lambda)$ respects the symmetries

$$(4.2) \quad M_{11}(\lambda) = \overline{M_{22}(\bar{\lambda})}, \quad M_{21}(\lambda) = -\overline{M_{12}(\bar{\lambda})}.$$

Moreover, if (4.2) holds, then only one of the two residue conditions (3.25a)-(3.25b) needs to be verified since the other condition is a consequence of symmetry. In view of (4.2), equation (4.1) can be written as

$$(4.3) \quad \begin{pmatrix} \overline{M_{22}(\bar{\lambda})} \\ -\overline{M_{12}(\bar{\lambda})} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j=1}^N \frac{C_j e^{2i\theta_j}}{\lambda - \lambda_j} \begin{pmatrix} M_{12}(\lambda_j) \\ M_{22}(\lambda_j) \end{pmatrix}.$$

Evaluation at $\bar{\lambda}_n$ yields

$$(4.4) \quad \begin{pmatrix} \overline{M_{22}(\lambda_n)} \\ -\overline{M_{12}(\lambda_n)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \sum_{j=1}^N \frac{C_j e^{2i\theta_j}}{\bar{\lambda}_n - \lambda_j} \begin{pmatrix} M_{12}(\lambda_j) \\ M_{22}(\lambda_j) \end{pmatrix}, \quad n = 1, \dots, N.$$

Solving this algebraic system for $M_{12}(\lambda_j)$ and $M_{22}(\lambda_j)$, $j = 1, \dots, N$, and substituting the resulting expression into (4.1) and (4.3) gives explicit expressions for the entries of $M(\lambda)$. The N -soliton solution $u(x, t)$ is then obtained from (3.23).

4.1 One-solitons

The one-solitons arise when $a(\lambda)$ has one purely imaginary zero λ_1 . Letting $N = 1$ in (4.4) and taking the complex conjugate of the second row, we find

$$\overline{M_{22}(\lambda_1)} = 1 + \frac{C_1 e^{2i\theta_1}}{\lambda_1 - \bar{\lambda}_1} M_{12}(\lambda_1), \quad -M_{12}(\lambda_1) = \frac{\bar{C}_1 e^{-2i\bar{\theta}_1}}{\lambda_1 - \bar{\lambda}_1} \overline{M_{22}(\lambda_1)}.$$

Solving these equations for $M_{12}(\lambda_1)$ and $\overline{M_{22}(\lambda_1)}$, and substituting the result into (4.3), we find

$$M_{12}(\lambda) = -\frac{4\bar{C}_1 \text{Im}(\lambda_1)^2}{(\lambda - \bar{\lambda}_1)(|C_1|^2 e^{2i\theta_1} + 4e^{2i\bar{\theta}_1} \text{Im}(\lambda_1)^2)}.$$

Equation (3.27) yields

$$Q_y = \frac{16\bar{C}_1 \text{Im}(\lambda_1)^2}{|C_1|^2 e^{2i\theta_1} + 4e^{2i\bar{\theta}_1} \text{Im}(\lambda_1)^2},$$

where $Q(y, t)$ is defined by (3.29). Letting $\lambda_1 = i\alpha$ and $C_1 = 2\alpha e^{x_0}$, where $\alpha > 0$ and $x_0 \in \mathbb{R}$ are parameters, we find a real-valued solution given by

$$(4.5) \quad Q(y, t) = 4 \arctan \left(e^{-2\alpha y - \frac{t}{2\alpha} + x_0} \right).$$

The solution $u(x, t)$ is given implicitly in terms of $Q(y, t)$ by (3.23). We recognize $Q(y, t)$ as a one-soliton solution of the sine-Gordon equation $Q_{ty} = \sin(Q)$. We will see in the next section that this is no coincidence.

5 Relation to sine-Gordon

The sine-Gordon equation

$$(5.1) \quad Q_{\tau y} = \sin(Q)$$

admits the Lax pair

$$(5.2) \quad \begin{cases} \phi_y + i\lambda\sigma_3\phi = -\frac{i}{2}Q_y\phi, \\ \phi_\tau - \frac{i}{4\lambda}\sigma_3\phi = \frac{i}{4\lambda}((\cos(Q) - 1)\sigma_3 - \sin(Q)\sigma_2)\phi, \end{cases}$$

where $\phi(y, \tau, \lambda)$ is a 2×2 -matrix valued eigenfunction. On the other hand, the change of variables $(x, t) \mapsto (y, \tau)$ with $y = p(x, t)$ and $\tau = t$ transforms the x -part of (3.6) into

$$\phi_y + i\lambda\sigma_3\phi = \frac{i}{2\sqrt{m}} \left(u_x + \frac{u_{xx}}{m} \right) \sigma_1 \phi.$$

This x -part coincides with that of (5.2) if we identify Q_y with

$$-\frac{1}{\sqrt{m}} \left(u_x + \frac{u_{xx}}{m} \right) = - \left(u_y + \frac{u_{yy}}{\sqrt{1 - u_y^2}} \right).$$

This is the identification made in (3.28) and suggests that if $u(x, t)$ satisfies (1.1), then the corresponding function $Q(y, \tau)$ evolves according to an equation in the sine-Gordon hierarchy. In fact, the following proposition shows that $Q(y, \tau)$ evolves exactly according to the sG equation.

Proposition 5.1 *Suppose $u(x, t)$ satisfies the generalized sG equation (1.1). Then the function $Q(y, \tau)$ defined by*

$$(5.3a) \quad Q(y, \tau) = -u(x, t) - \arcsin \left(\frac{u_x(x, t)}{\sqrt{1 + u_x^2(x, t)}} \right),$$

$$(5.3b) \quad y = x - t + \int_{-\infty}^x (\sqrt{1 + u_x^2(x', t)} - 1) dx', \quad \tau = t,$$

satisfies the sG equation (5.1).

Proof. In view of the conservation law (3.8), we have the relations

$$(5.4) \quad \frac{\partial}{\partial t} = \frac{\partial}{\partial \tau} - \cos(u)\sqrt{m}\frac{\partial}{\partial y}, \quad \frac{\partial}{\partial x} = \sqrt{m}\frac{\partial}{\partial y}.$$

Using these relations we find that the generalized sG equation (1.1) in terms of the variables (y, τ) takes the form

$$(5.5) \quad u_{y\tau} = \sqrt{1 - u_y^2} \sin(u).$$

Assume that $u(x, t)$ satisfies (1.1). Applying $\partial_\tau \partial_y$ to (5.3a) and using (5.5) and its y -derivative to replace $u_{yy\tau}$ and $u_{y\tau}$ in the resulting equation, we find

$$Q_{y\tau} = -u_y \cos(u) - \sqrt{1 - u_y^2} \sin(u).$$

Replacing u with $-Q - \arcsin(u_y)$ in this equation and simplifying, we find that the right-hand side equals $\sin(Q)$. \square

6 Conservation laws

An infinite sequence of conservation laws can be derived for a bi-Hamiltonian equation by means of the Lenard scheme. This approach however involves the inversion of the bi-Hamiltonian operators. In the case of equation (1.1) the operators θ_1 and θ_2 are nonlocal and difficult to invert. Instead of following this procedure we can derive conservation laws of (1.1) by utilizing the connection with the sG equation established in the previous section. Indeed, suppose that the sG equation (5.1) admits a conservation law of the form

$$(6.1) \quad A_\tau = B_y,$$

where A and B are functionals of $Q(y, \tau)$ and its y -derivatives. The change of variables (5.3b) transforms (6.1) into

$$(6.2) \quad (\sqrt{m}A)_t = (B - \cos(u)\sqrt{m}A)_x.$$

Writing this equation in terms of $u(x, t)$ by means of relation (5.3a), we find a conservation law for equation (1.1). Since the conservation laws of sG can be constructed explicitly, this approach yields an infinite number of conservation laws for (1.1). We next illustrate this method by deriving a few conservation laws explicitly.

As noted in section 2, equation (5.1) admits the recursion operator $J_2 J_1^{-1}$ where the Hamiltonian operators are

$$J_1 = \partial_y, \quad J_2 = \partial_y^3 + \partial_y q \partial_y^{-1} q \partial_y,$$

and $q = Q_y$. The sG equation is given by

$$q_\tau = J_1 \frac{\delta H_0}{\delta q},$$

where the Hamiltonian H_0 is defined by

$$H_0 = \int (\cos(\partial_y^{-1} q) - 1) dy.$$

Further conservation laws can be determined according to the equation

$$\frac{\delta H_{n+1}}{\delta q} = J_1^{-1} J_2 \frac{\delta H_n}{\delta q}.$$

Since J_1^{-1} is given by simple integration, the conserved quantities H_n for $n \geq 0$ can be constructed recursively. We let $H_n = \int A_n dy$ and write the associated conservation law in the form

$$H_n : (A_n)_\tau = (B_n)_y.$$

We find the following explicit expressions for the first few conservation laws expressed in terms of Q :

$$\begin{aligned} H_0 : (\cos Q - 1)_\tau &= \left(-\frac{1}{2}(\partial_y^{-1} \sin Q)^2 \right)_y, \\ H_1 : \left(\frac{1}{2}Q_y^2 \right)_\tau &= (-\cos Q)_y, \\ H_2 : \left(\frac{1}{4}Q_y^4 - Q_{yy}^2 \right)_\tau &= (-Q_y^2 \cos Q)_y, \\ H_3 : \left(\frac{1}{8}Q_y^6 - \frac{5}{2}Q_y^2 Q_{yy}^2 + Q_{yyy}^2 \right)_\tau &= \left(-\frac{3}{4}Q_y^4 \cos(Q) - 2Q_y^2 Q_{yy} \sin(Q) + Q_{yy}^2 \cos(Q) \right)_y. \end{aligned}$$

We use (5.3a), (5.4), and (6.2) to rewrite these conservation laws in terms of $u(x, t)$. The result for H_0 and H_1 is

$$\begin{aligned} H_0 : & (\cos(u) + (\cos u)_x - \sqrt{m})_t \\ &= \left(-\frac{1}{2}(\partial_x^{-1}(\sin u + (\sin u)_x))^2 + (\cos(u) + (\cos u)_x - \sqrt{m}) \cos u \right)_x, \\ H_1 : & \left(\frac{(mu_x + u_{xx})^2}{2m^{5/2}} \right)_t \\ &= \left(\frac{(-m(m^2 + m + m_x) - u_{xx}^2) \cos(u(x, t)) + 2m^2 u_x \sin(u(x, t))}{2m^{5/2}} \right)_x. \end{aligned}$$

For H_2 and H_3 the expressions are quite long, thus we only present the conserved quantities:

$$\begin{aligned} H_2 &= \int \frac{1}{8m^{3/2}} \left[\left(\frac{u_{xx}}{m} + u_x \right)^4 - \frac{(3m_x u_{xx} - 2m(u_{xx} + u_{xxx}))^2}{m^4} \right] dx, \\ H_3 &= \int \frac{1}{16m^{5/2}} \left[-\frac{5(mu_x + u_{xx})^2(3m_x u_{xx} - 2m(u_{xx} + u_{xxx}))^2}{m^6} + \left(\frac{u_{xx}}{m} + u_x \right)^6 \right. \\ &\quad \left. + \frac{8(m^2(u_{xxx} + u_{xxxx}) + m(-2m_x u_{xx} - 5m_x u_{xxx} + 15u_{xx}^3) - 18u_{xx}^3)^2}{m^6} \right] dx. \end{aligned}$$

It can be verified directly that these indeed are conservation laws for equation (1.1).

7 Traveling-wave solutions

For a traveling wave $u(x, t) = \varphi(x - ct)$ traveling with speed c , equation (1.1) takes the form

$$(7.1) \quad (\cos(\varphi) - c)\varphi_{xx} = \sin(\varphi) + \varphi_x^2 \sin(\varphi).$$

We rewrite (7.1) as

$$\frac{\varphi_{xx}}{1 + \varphi_x^2} = \frac{\sin(\varphi)}{\cos(\varphi) - c}.$$

Multiplying this equation by φ_x and integrating with respect to x , we find

$$(7.2) \quad \frac{1}{2} \log(1 + \varphi_x^2) = -\log|c - \cos(\varphi)| + \log(b),$$

where $b > 0$ is an integration constant. Exponentiating and then taking the square of both sides of equation (7.2), we find after rearrangement

$$(7.3) \quad \varphi_x^2 = F(\varphi), \quad F(\varphi) = \frac{b^2 - (c - \cos(\varphi))^2}{(c - \cos(\varphi))^2}.$$

The qualitative structure of the traveling-wave solutions of (1.1) can be obtained by analyzing this ordinary differential equation. It is in fact possible to solve (7.3) explicitly, which yields an expression for φ^{-1} in terms of elliptic functions. We proceed instead by briefly describing the qualitative structure of a few different types of traveling waves. We refer to [14] for a more detailed analysis in the similar case of the Camassa-Holm equation.

The qualitative structure of solutions of (7.3) is determined by the zeros and the poles of $F(\varphi)$. The sets of zeros and poles of $F(\varphi)$ are given by

$$\{\varphi | \cos(\varphi) = c \pm b\} \quad \text{and} \quad \{\varphi | \cos(\varphi) = c\},$$

respectively. Since $b > 0$, these sets are disjoint.

The traveling waves are parametrized by $b > 0$ and $c \in \mathbb{R}$. Since we are interested in waves such that $\lim_{x \rightarrow \pm\infty} \varphi(x) \in 2\pi\mathbb{Z}$ we assume that $b = |c - 1|$. In this case $F(\varphi)$ has double zeros at $\varphi = 2n\pi$, $n \in \mathbb{Z}$, and these zeros give rise to solutions $\varphi(x)$ of (7.3) which approach elements in the set $2\pi\mathbb{Z}$ exponentially fast as $x \rightarrow \pm\infty$. In addition to these solutions there exists a range of periodic solutions obtained for values of b such that $b \neq |c - 1|$.

Note that if $u(x, t)$ solves equation (1.1), then so do the functions $(x, t) \mapsto \pm u(x, t) + 2\pi n$, $n \in \mathbb{Z}$. This corresponds to the fact that $F(\varphi)$ is an even function of period 2π .

7.1 Smooth waves

Assume that $b = |c - 1|$ and $c < -1$. In this case $F(\varphi)$ has double zeros at $\varphi = 2n\pi$, $n \in \mathbb{Z}$, and $F(\varphi) > 0$ away from these zeros, see Figure 1. Moreover, the set of poles of $F(\varphi)$ is empty. An analysis of (7.3) shows that corresponding to the interval between two consecutive zeros $2n\pi$ and $2(n + 1)\pi$ there exist two smooth solutions $\varphi_1(x)$ and $\varphi_2(x)$ such that

$$\lim_{x \rightarrow -\infty} \varphi_1(x) = 2n\pi, \quad \lim_{x \rightarrow \infty} \varphi_1(x) = 2(n + 1)\pi.$$

and

$$\lim_{x \rightarrow -\infty} \varphi_2(x) = 2(n + 1)\pi, \quad \lim_{x \rightarrow \infty} \varphi_2(x) = 2n\pi.$$

This leads to the following result.

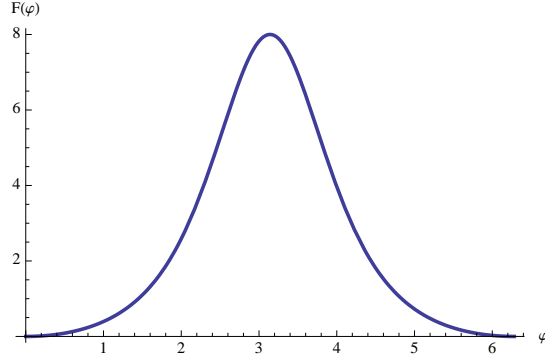


Figure 1 One period of the function $F(\varphi)$ for $b = |c - 1|$ and $c = -2$.

Proposition 7.1 *For any $c < -1$ and $n \in \mathbb{Z}$, there exists:*

- A smooth traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow -\infty} \varphi(x) = 2n\pi, \quad \lim_{x \rightarrow \infty} \varphi(x) = 2(n+1)\pi.$$

- A smooth traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow -\infty} \varphi(x) = 2(n+1)\pi, \quad \lim_{x \rightarrow \infty} \varphi(x) = 2n\pi.$$

The profiles of these waves for $n = 0$ and $c = -2$ are shown in (a) and (b) of Figure 3.

7.2 Cusped waves

In addition to the above smooth traveling waves, equation (1.1) also admits cusped traveling waves (cuspons). These arise for $b = |c - 1|$ and $-1 \leq c < 1$ (if $c \geq 1$, then $F(\varphi) \leq 0$ for all $\varphi \in \mathbb{R}$, and there do not exist nonconstant solutions of (7.3)).

Assume first that $-1 < c < 0$. In this case $F(\varphi)$ has double zeros at $\varphi = 2n\pi$, $n \in \mathbb{Z}$ and $F(\varphi) > 0$ away from these zeros. See (b) of Figure 2 for the graph of the numerator of $F(\varphi)$ when $c = -1/10$. Moreover, $F(\varphi)$ has double poles at $\varphi = \pm \arccos(c) + 2n\pi$, $n \in \mathbb{Z}$. Near one of the poles $\varphi_0 = \pm \arccos(c) + 2n\pi$, the behavior of $F(\varphi)$ is given by

$$(7.4) \quad F(\varphi) = \frac{1-c}{(1+c)(\varphi - \varphi_0)^2} + O\left(\frac{1}{\varphi - \varphi_0}\right), \quad \varphi \rightarrow \varphi_0.$$

Let $\varphi(x)$ be a solution of (7.3) such that $\varphi(x_0)$ belongs to the interval $(0, \arccos(c))$ and $\varphi_x(x_0) > 0$ at some point x_0 . Then, as x increases above x_0 , the value of $\varphi(x)$ approaches $\varphi_0 = \arccos(c)$. At some point $x_c > x_0$, we have $\varphi(x_c) = \varphi_0$ and near this point integration of (7.3) using (7.4) shows that the behavior of $\varphi(x)$ is given by

$$(7.5) \quad |\varphi(x) - \varphi_0| = \sqrt{2} \left(\frac{1-c}{1+c} \right)^{1/4} \sqrt{|x - x_c|} + O(x - x_c), \quad x \rightarrow x_c.$$

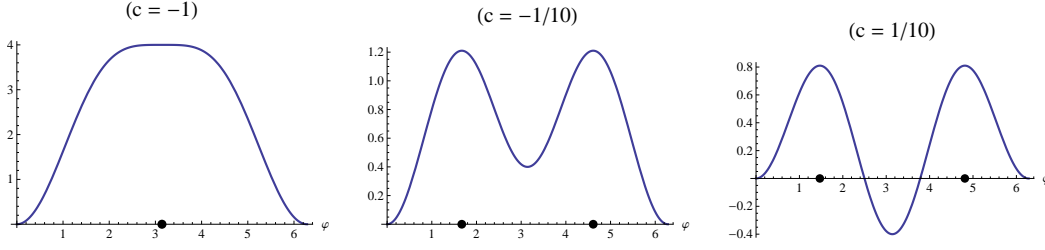


Figure 2 The numerator $b^2 - (c - \cos(\varphi))^2$ of the function $F(\varphi)$ for $b = |c - 1|$ and $c = -1$ (left), $c = -1/10$ (middle), $c = 1/10$ (right). The black dots on the φ -axes denote the values of φ at which $F(\varphi)$ has poles.

In the case when $\varphi(x) - \varphi_0$ has the same sign locally for $x < x_c$ and $x > x_c$, we say that $\varphi(x)$ has a *cusp* at $x = x_c$, i.e. $\varphi(x)$ is continuous, smooth locally on both sides of x_c , and

$$\lim_{x \uparrow x_c} \varphi_x(x) = - \lim_{x \downarrow x_c} \varphi_x(x) = \pm \infty.$$

Cusps such that $\varphi(x)$ has a minimum at x_c are also called *anticusps*. As $x \rightarrow \pm\infty$, $\varphi(x)$ decays exponentially to 0. A similar analysis applies to the interval $(-\arccos(c), 0)$ and the translations of these intervals by multiples of 2π . This proves the following result in the case of $-1 < c < 0$:

Proposition 7.2 *For any $-1 < c < 1$, $c \neq 0$, and $n \in \mathbb{Z}$, there exists:*

- A cusped traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow \pm\infty} \varphi(x) = 2n\pi, \quad \max_{x \in \mathbb{R}} \varphi(x) = 2n\pi + \arccos(c), \quad \inf_{x \in \mathbb{R}} \varphi(x) = 2n\pi.$$

The maximum is attained at some point x_c , and near x_c the behavior of $\varphi(x)$ is given by (7.5) with $\varphi_0 = 2n\pi + \arccos(c)$, see (c) of Figure 3.

- An anticusped traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow \pm\infty} \varphi(x) = 2n\pi, \quad \min_{x \in \mathbb{R}} \varphi(x) = 2n\pi - \arccos(c), \quad \sup_{x \in \mathbb{R}} \varphi(x) = 2n\pi.$$

The minimum is attained at some point x_c , and near x_c the behavior of $\varphi(x)$ is given by (7.5) with $\varphi_0 = 2n\pi - \arccos(c)$, see (d) of Figure 3.

For $c = -1$ and any $n \in \mathbb{Z}$, there exists:

- A cusped traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow \pm\infty} \varphi(x) = 2n\pi, \quad \max_{x \in \mathbb{R}} \varphi(x) = (2n + 1)\pi, \quad \inf_{x \in \mathbb{R}} \varphi(x) = 2n\pi.$$

The maximum is attained at some point x_c , and near x_c the behavior of $\varphi(x)$ is given by

$$(7.6) \quad |\varphi(x) - \varphi_0| = 12^{1/3} |x - x_c|^{1/3} + O(|x - x_c|^{2/3}), \quad x \rightarrow x_c.$$

with $\varphi_0 = (2n + 1)\pi$, see (e) of Figure 3.

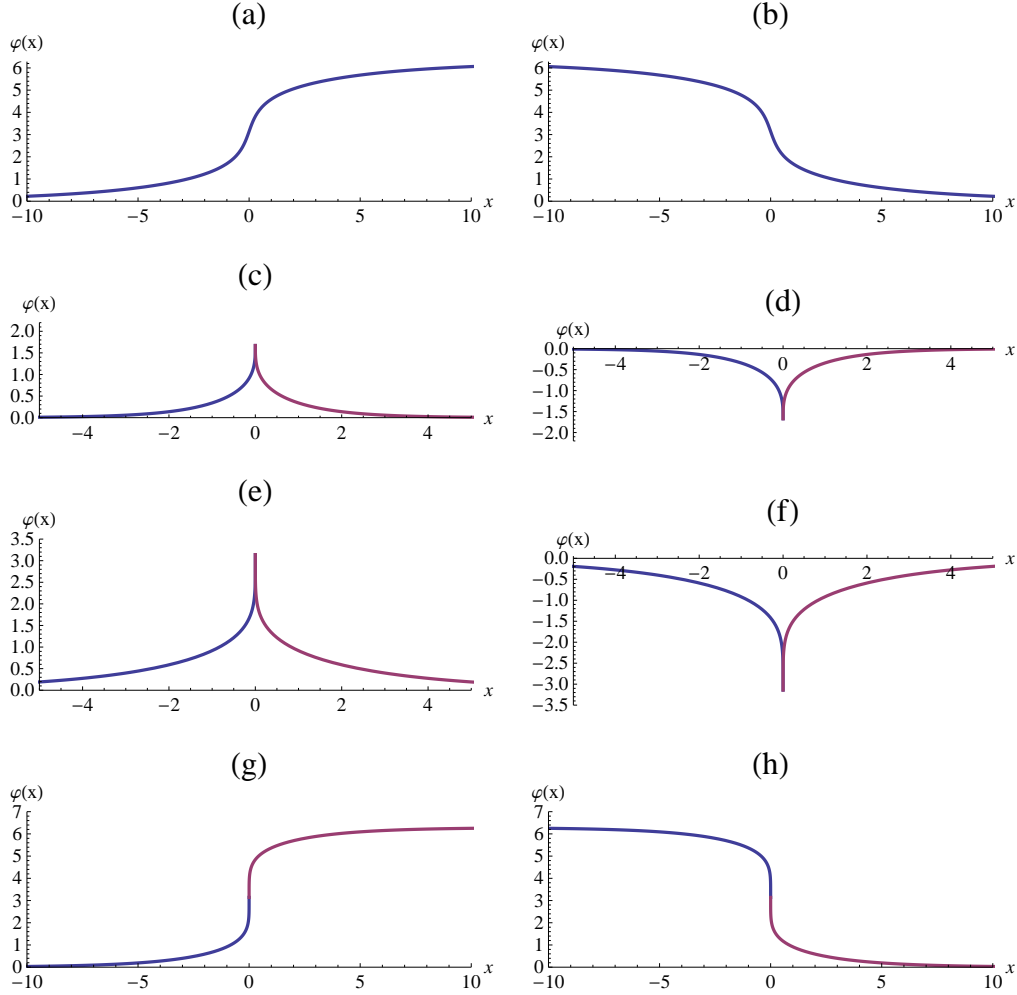


Figure 3 Profiles of different types of traveling-wave solutions $u(x, t) = \varphi(x - ct)$ of equation (1.1): (a), (b) Smooth traveling waves with $c = -2$; (c) Cusped traveling wave with $c = -1/10$; (d) Anticusped traveling wave with $c = -1/10$; (e) Cusped traveling wave with $c = -1$; (f) Anticusped traveling wave with $c = -1$; (g), (h) Traveling waves with $c = -1$ which combine the cusped and anticusped solutions.

- An anticusped traveling-wave solution $u(x, t) = \varphi(x - ct)$ of (1.1) such that

$$\lim_{x \rightarrow \pm\infty} \varphi(x) = 2n\pi, \quad \min_{x \in \mathbb{R}} \varphi(x) = (2n - 1)\pi, \quad \sup_{x \in \mathbb{R}} \varphi(x) = 2n\pi.$$

The minimum is attained at some point x_c , and near x_c the behavior of $\varphi(x)$ is given by (7.6) with $\varphi_0 = (2n - 1)\pi$, see (f) of Figure 3.

Proof. It remains to prove the cases $c = -1$ and $0 < c < 1$. For $0 < c < 1$, $F(\varphi)$ has double zeros at $\varphi = 2n\pi$, $n \in \mathbb{Z}$, simple zeros at $\varphi = \pm \arccos(2c - 1) + 2n\pi$, $n \in \mathbb{Z}$, and double poles at $\varphi = \pm \arccos(c) + 2n\pi$, $n \in \mathbb{Z}$, see Figure 2. Since $2c - 1 < c$ for $0 < c < 1$, we have $\arccos(c) < \arccos(2c - 1)$. Thus, $F(\varphi)$ behaves qualitatively in the interval $[0, 2\pi]$ as follows:

1. $F(\varphi) > 0$ in the interval $(0, \arccos(c))$. $F(\varphi)$ has a double zero at $\varphi = 0$ and a double pole at $\varphi = \arccos(c)$.
2. $F(\varphi) > 0$ in the interval $(\arccos(c), \arccos(2c - 1))$. $F(\varphi)$ has a simple zero at $\varphi = \arccos(2c - 1)$.
3. $F(\varphi) < 0$ in the interval $(\arccos(2c - 1), 2\pi - \arccos(2c - 1))$.
4. $F(\varphi) > 0$ in the interval $(2\pi - \arccos(2c - 1), 2\pi - \arccos(c))$. $F(\varphi)$ has a simple zero at $\varphi = 2\pi - \arccos(2c - 1)$ and a double pole at $\varphi = 2\pi - \arccos(c)$.
5. $F(\varphi) > 0$ in the interval $(2\pi - \arccos(c), 2\pi)$. $F(\varphi)$ has a double zero at $\varphi = 2\pi$.

The proof when $0 < c < 1$ follows from the above properties of the function $F(\varphi)$ and from its 2π periodicity. Indeed, the solutions listed in the proposition arise in the intervals $(0, \arccos(c))$ and $(2\pi - \arccos(c), 2\pi)$, as well as in the 2π -translations of these intervals. The intervals $(\arccos(c), \arccos(2c - 1))$ and $(2\pi - \arccos(2c - 1), 2\pi - \arccos(c))$ give rise to *periodic* cusped traveling waves, which are not included in the proposition.

For $c = -1$, $F(\varphi)$ has double zeros at $\varphi = 2n\pi$, $n \in \mathbb{Z}$, $F(\varphi) > 0$ away from these zeros, and $F(\varphi)$ has fourth-order poles at $(2n + 1)\pi$, $n \in \mathbb{Z}$. The behavior of $F(\varphi)$ near one of the poles $\varphi_0 = (2n + 1)\pi$ is given by

$$(7.7) \quad F(\varphi) = \frac{16}{(\varphi - \varphi_0)^4} + O\left(\frac{1}{(\varphi - \varphi_0)^3}\right), \quad \varphi \rightarrow \varphi_0.$$

It follows that the behavior of a solution $\varphi(x)$ of (7.3) near a point x_c at which $\varphi(x_c) = \varphi_0$ is given by (7.6). The rest of the proof when $c = -1$ follows as in the previous cases. \square

Since both the cusped and anticusped waves of Proposition 7.2 when $c = -1$ satisfy $\varphi \in (2\mathbb{Z} + 1)\pi$ at the cusps, it is possible to combine these waves into traveling waves such as those shown in (g) and (h) of Figure 3. These waves are reminiscent of the smooth traveling-wave solutions but have infinite slope at the points where $\varphi \in \pi + 2\pi\mathbb{Z}$.

We emphasize that we have only shown formally that the cusped traveling waves are solutions of (1.1). In particular, we have not provided an appropriate weak formulation of (1.1) of which these waves are solutions. However, in analogy with the CH equation and the form of Eq. (7.1), it is expected that such a weak formulation does exist. Indeed, let $\varphi(x)$ be a traveling wave with a cusp located at $x = x_c$. Then φ_{xx} will have a singularity at x_c . However, since φ_{xx} in equation (7.1) is multiplied by the factor $\cos(\varphi) - c$ which vanishes at x_c , it is possible that the product $\varphi_{xx}(\cos(\varphi) - c)$ is sufficiently regular to allow for a weak formulation. Actually, a direct analog of this mechanism is responsible for the cancellation of the singularities for the peaked and cusped solutions of the Camassa-Holm equation. If this analogy with CH turns out to be correct, it should also be possible to construct more exotic traveling-wave solutions of equation (1.1), such as fractal-like waves and waves with plateaus cf. [14].

We have not investigated the stability of the cusped traveling waves of (1.1). It is in fact unknown even in the case of the Camassa-Holm and Degasperis-Procesi equations whether the cusped solutions are stable (although the peaked solutions of these equations are known to be stable [5, 13, 15]).

Acknowledgement *The authors are grateful to Sergei Sakovich for bringing the references [2, 16, 17] to their attention. ASF acknowledges the support of the Guggenheim foundation, USA.*

References

- [1] M. J. Ablowitz, D. J. Kaup, A. C. Newell, and H. Segur, Method for solving the sine-Gordon equation, *Phys. Rev. Lett.* **30** (1973), 1262–1264.
- [2] R. Beals, M. Rabelo, and K. Tenenblat, Bäcklund transformations and inverse scattering solutions for some pseudospherical surface equations, *Stud. Appl. Math.* **81** (1989), 125–151.
- [3] R. Camassa and D. D. Holm, An integrable shallow water equation with peaked solitons, *Phys. Rev. Lett.* **71** (1993), 1661–1664.
- [4] A. Constantin, V. S. Gerdjikov, and R. I. Ivanov, Inverse scattering transform for the Camassa-Holm equation, *Inverse Problems* **22** (2006), 2197–2207.
- [5] A. Constantin and W. Strauss, Stability of peakons, *Comm. Pure Appl. Math.* **53** (2000), 603–610.
- [6] A. S. Fokas, On a class of physically important integrable equations, *Phys. D* **87** (1995), 145–150.
- [7] A. S. Fokas, A unified transform method for solving linear and certain nonlinear PDEs, *Proc. Roy. Soc. Lond. A* **453** (1997), 1411–1443.
- [8] A. S. Fokas and R. L. Anderson, On the use of isospectral eigenvalue problems for obtaining hereditary symmetries for Hamiltonian systems, *J. Math. Phys.* **23** (1982), 1066–1073.

- [9] A. S. Fokas and A. R. Its, The linearization of the initial-boundary value problem of the nonlinear Schrödinger equation, *SIAM J. Math. Anal.* **27** (1996), 738–764.
- [10] A. S. Fokas and J. Lenells, Explicit soliton asymptotics for the Korteweg-de Vries equation on the half-line, preprint, arXiv:0812.1579.
- [11] B. Fuchssteiner and A. S. Fokas, Symplectic structures, their Bäcklund transformation and hereditary symmetries, *Physica D* **4** (1981), 47–66.
- [12] J. Lenells, The scattering approach for the Camassa-Holm equation, *J. Nonlinear Math. Phys.* **9** (2002), 389–393.
- [13] J. Lenells, Stability of periodic peakons, *Internat. Math. Res. Notices* **2004** (2004), 485–499.
- [14] J. Lenells, Traveling wave solutions of the Camassa-Holm equation, *J. Diff. Eq.* **217** (2005), 393–430.
- [15] Z. Lin and Y. Liu, Stability of peakons for the Degasperis-Procesi equation, *Comm. Pure Appl. Math.* **62** (2009), 125–146.
- [16] M. L. Rabelo, On equations which describe pseudospherical surfaces, *Stud. Appl. Math.* **81** (1989), 221–248.
- [17] A. Sakovich and S. Sakovich, On transformations of the Rabelo equations, *SIGMA Symmetry Integrability Geom. Methods Appl.* **3** (2007), Paper 086, 8 pp.